



OIKE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/888,320

DATE: 02/21/2002
TIME: 12:01:24

Input Set : A:\Nih413-1.app
Output Set: N:\CRF3\02212002\I888320.raw

Does Not Comply
Corrected Diskette Needed
x and

```

3 <110> APPLICANT: Barry III, Clifton E.
4   DeBarber, Andrea E.
5   Mdluli, Khisimuzi
6   Bekker, Linda-Gail
7   The Government of the United States of America
8   as represented by The Secretary of the
9   Department of Health and Human Services
11 <120> TITLE OF INVENTION: Methods of Diagnosing Multidrug Resistant Tuberculosis
13 <130> FILE REFERENCE: 015280-413100US
15 <140> CURRENT APPLICATION NUMBER: US 09/888,320
16 <141> CURRENT FILING DATE: 2001-06-22
18 <150> PRIOR APPLICATION NUMBER: US 60/214,187
19 <151> PRIOR FILING DATE: 2000-06-26
21 <160> NUMBER OF SEQ ID NOS: 16
23 <170> SOFTWARE: PatentIn Ver. 2.1
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 1867
27 <212> TYPE: DNA
28 <213> ORGANISM: Mycobacterium tuberculosis
30 <220> FEATURE:
31 <223> OTHER INFORMATION: wild-type EtaA monooxygenase (Rv3854c, EthA)
33 <220> FEATURE:
34 <221> NAME/KEY: CDS
35 <222> LOCATION: (201)..(1670)
36 <223> OTHER INFORMATION: EtaA
38 <400> SEQUENCE: 1
39 agcggacggt cctcgagaag gttctcggcg gtggcgagga tcgccagttc acgatcgtcg 60
41 ccggacggcc gcgcggtgcg ccggccccta ggcagcgaag cctgactggc cgcggaggtg 120
43 gtcaccctgg cagcttacta cgtgtcgata gtgtcgacat ctggttgacg gcctcgacat 180
45 tacgttgata gcgtggatcc atg acc gag cac ctg gac gtt gtc atc gtg ggc 233
46                               Met Thr Glu His Leu Asp Val Val Ile Val Gly
47                               1           5           10
49 gct gga atc tcc ggt gtc agc gcg gcc tgg cac ctg cag gac cgt tgc 281
50 Ala Gly Ile Ser Gly Val Ser Ala Ala Trp His Leu Gln Asp Arg Cys
51                               15           20           25
53 ccg acc aag agc tac gcc atc ctg gaa aag cgg gaa tcc atg ggc ggc 329
54 Pro Thr Lys Ser Tyr Ala Ile Leu Glu Lys Arg Glu Ser Met Gly Gly
55                               30           35           40
57 acc tgg gat ttg ttc cgt tat ccc gga att cgc tcc gac tcc gac atg 377
58 Thr Trp Asp Leu Phe Arg Tyr Pro Gly Ile Arg Ser Asp Ser Asp Met
59                               45           50           55
61 tac acg cta ggt ttc cga ttc cgt ccc tgg acc gga cgg cag gcg atc 425
62 Tyr Thr Leu Gly Phe Arg Phe Arg Pro Trp Thr Gly Arg Gln Ala Ile

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/888,320

DATE: 02/21/2002

TIME: 12:01:24

Input Set : A:\Nih413-1.app

Output Set: N:\CRF3\02212002\I888320.raw

63	60		65		70		75										
65	gcc	gac	ggc	aag	ccc	atc	ctc	gag	tac	gtc	aag	agc	acc	gcg	gcc	atg	473
66	Ala	Asp	Gly	Lys	Pro	Ile	Leu	Glu	Tyr	Val	Lys	Ser	Thr	Ala	Ala	Met	
67				80						85					90		
69	tat	gga	atc	gac	agg	cat	atc	cgg	ttc	cac	cac	aag	gtg	atc	agt	gcc	521
70	Tyr	Gly	Ile	Asp	Arg	His	Ile	Arg	Phe	His	His	Lys	Val	Ile	Ser	Ala	
71				95					100					105			
73	gat	tgg	tcg	acc	gcg	gaa	aac	cgc	tgg	acc	gtt	cac	atc	caa	agc	cac	569
74	Asp	Trp	Ser	Thr	Ala	Glu	Asn	Arg	Trp	Thr	Val	His	Ile	Gln	Ser	His	
75			110					115					120				
77	ggc	acg	ctc	agc	gcc	ctc	acc	tgc	gaa	ttc	ctc	ttt	ctg	tgc	agc	ggc	617
78	Gly	Thr	Leu	Ser	Ala	Leu	Thr	Cys	Glu	Phe	Leu	Phe	Leu	Cys	Ser	Gly	
79			125					130					135				
81	tac	tac	aac	tac	gac	gag	ggc	tac	tcg	ccg	aga	ttc	gcc	ggc	tcg	gag	665
82	Tyr	Tyr	Asn	Tyr	Asp	Glu	Gly	Tyr	Ser	Pro	Arg	Phe	Ala	Gly	Ser	Glu	
83	140					145					150				155		
85	gat	ttc	gtc	ggg	ccg	atc	atc	cat	ccg	cag	cac	tgg	ccc	gag	gac	ctc	713
86	Asp	Phe	Val	Gly	Pro	Ile	Ile	His	Pro	Gln	His	Trp	Pro	Glu	Asp	Leu	
87				160						165				170			
89	gac	tac	gac	gct	aag	aac	atc	gtc	gtg	atc	ggc	agt	ggc	gca	acg	gcg	761
90	Asp	Tyr	Asp	Ala	Lys	Asn	Ile	Val	Val	Ile	Gly	Ser	Gly	Ala	Thr	Ala	
91			175						180					185			
93	gtc	acg	ctc	gtg	ccg	gcg	ctg	gcg	gac	tcg	ggc	gcc	aag	cac	gtc	acg	809
94	Val	Thr	Leu	Val	Pro	Ala	Leu	Ala	Asp	Ser	Gly	Ala	Lys	His	Val	Thr	
95			190						195					200			
97	atg	ctg	cag	cgc	tca	ccc	acc	tac	atc	gtg	tcg	cag	cca	gac	cgg	gac	857
98	Met	Leu	Gln	Arg	Ser	Pro	Thr	Tyr	Ile	Val	Ser	Gln	Pro	Asp	Arg	Asp	
99		205					210						215				
101	ggc	atc	gcc	gag	aag	ctc	aac	cgc	tgg	ctg	ccg	gag	acc	atg	gcc	tac	905
102	Gly	Ile	Ala	Glu	Lys	Leu	Asn	Arg	Trp	Leu	Pro	Glu	Thr	Met	Ala	Tyr	
103	220					225					230				235		
105	acc	gcg	gta	cgg	tgg	aag	aac	gtg	ctg	cgc	cag	gcg	gcc	gtg	tac	agc	953
106	Thr	Ala	Val	Arg	Trp	Lys	Asn	Val	Leu	Arg	Gln	Ala	Ala	Val	Tyr	Ser	
107				240						245				250			
109	gcc	tgc	cag	aag	tgg	cca	cgg	cgc	atg	cgg	aag	atg	ttc	ctg	agc	ctg	1001
110	Ala	Cys	Gln	Lys	Trp	Pro	Arg	Arg	Met	Arg	Lys	Met	Phe	Leu	Ser	Leu	
111			255						260					265			
113	atc	cag	cgc	cag	cta	ccc	gag	ggg	tac	gac	gtg	cga	aag	cac	ttc	ggc	1049
114	Ile	Gln	Arg	Gln	Leu	Pro	Glu	Gly	Tyr	Asp	Val	Arg	Lys	His	Phe	Gly	
115			270						275					280			
117	ccg	cac	tac	aac	ccc	tgg	gac	cag	cga	ttg	tgc	ttg	gtg	ccc	aac	ggc	1097
118	Pro	His	Tyr	Asn	Pro	Trp	Asp	Gln	Arg	Leu	Cys	Leu	Val	Pro	Asn	Gly	
119		285					290						295				
121	gac	ctg	ttc	cgg	gcc	att	cgt	cac	ggg	aag	gtc	gag	gtg	gtg	acc	gac	1145
122	Asp	Leu	Phe	Arg	Ala	Ile	Arg	His	Gly	Lys	Val	Glu	Val	Val	Thr	Asp	
123	300					305					310				315		
125	acc	att	gaa	cgg	ttc	acc	gcg	acc	gga	atc	cgg	ctg	aac	tca	ggt	cgc	1193
126	Thr	Ile	Glu	Arg	Phe	Thr	Ala	Thr	Gly	Ile	Arg	Leu	Asn	Ser	Gly	Arg	
127				320						325				330			

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/888,320

DATE: 02/21/2002

TIME: 12:01:24

Input Set : A:\Nih413-1.app

Output Set: N:\CRF3\02212002\I888320.raw

```

129 gaa ctg ccg gct gac atc atc att acc gca acg ggg ttg aac ctg cag 1241
130 Glu Leu Pro Ala Asp Ile Ile Ile Thr Ala Thr Gly Leu Asn Leu Gln
131 335 340 345
133 ctt ttt ggt ggg gcg acg gcg act atc gac gga caa caa gtg gac atc 1289
134 Leu Phe Gly Gly Ala Thr Ala Thr Ile Asp Gly Gln Gln Val Asp Ile
135 350 355 360
137 acc acg acg atg gcc tac aag ggc atg atg ctt tcc ggc atc ccc aac 1337
138 Thr Thr Thr Met Ala Tyr Lys Gly Met Met Leu Ser Gly Ile Pro Asn
139 365 370 375
141 atg gcc tac acg gtt ggc tac acc aat gcc tcc tgg acg ctg aag gcc 1385
142 Met Ala Tyr Thr Val Gly Tyr Thr Asn Ala Ser Trp Thr Leu Lys Ala
143 380 385 390 395
145 gac ctg gtg tcg gag ttt gtc tgt cgc ttg ttg aat tac atg gac gac 1433
146 Asp Leu Val Ser Glu Phe Val Cys Arg Leu Leu Asn Tyr Met Asp Asp
147 400 405 410
149 aac ggt ttt gac acc gtg gtc gtc gag cga ccg ggc tca gat gtc gaa 1481
150 Asn Gly Phe Asp Thr Val Val Val Glu Arg Pro Gly Ser Asp Val Glu
151 415 420 425
153 gag cgg ccc ttc atg gag ttc acc cca ggt tac gtg ctg cgc tcg ctg 1529
154 Glu Arg Pro Phe Met Glu Phe Thr Pro Gly Tyr Val Leu Arg Ser Leu
155 430 435 440
157 gac gag ctg ccc aag cag ggt tcg cgt aca ccg tgg cgc ctg aat cag 1577
158 Asp Glu Leu Pro Lys Gln Gly Ser Arg Thr Pro Trp Arg Leu Asn Gln
159 445 450 455
161 aac tac cta cgt gac atc cgg ctc atc cgg cgc ggc aag atc gac gac 1625
162 Asn Tyr Leu Arg Asp Ile Arg Leu Ile Arg Arg Gly Lys Ile Asp Asp
163 460 465 470 475
165 gag ggt ctg cgg ttc gcc aaa agg cct gcc ccg gtg ggg gtt tag 1670
166 Glu Gly Leu Arg Phe Ala Lys Arg Pro Ala Pro Val Gly Val
W--> 167 480 485 490
169 ctttagcgac ggttttagcgc cggtttaggc catagtcaga cgacgatgat gccgtogtcg 1730
171 tcgctgtagg cgatatcgcc cggaacgaat gtcaccccgcc ccagcgtgat ttcaacgtcg 1790
173 cgttctccgg caccggtctt ggtgctcttg cggggatttg tgcccagcgc tttgatgccg 1850
175 atgtcgatgc cgcgcag 1867
178 <210> SEQ ID NO: 2
179 <211> LENGTH: 489
180 <212> TYPE: PRT
181 <213> ORGANISM: Mycobacterium tuberculosis > 22257
182 <223> OTHER INFORMATION: wild-type EtaA monooxygenase (Rv3854c, EthA)
184 <400> SEQUENCE: 2
185 Met Thr Glu His Leu Asp Val Val Ile Val Gly Ala Gly Ile Ser Gly
186 1 5 10 15
187 Val Ser Ala Ala Trp His Leu Gln Asp Arg Cys Pro Thr Lys Ser Tyr
188 20 25 30
189 Ala Ile Leu Glu Lys Arg Glu Ser Met Gly Gly Thr Trp Asp Leu Phe
190 35 40 45
191 Arg Tyr Pro Gly Ile Arg Ser Asp Ser Asp Met Tyr Thr Leu Gly Phe
192 50 55 60
193 Arg Phe Arg Pro Trp Thr Gly Arg Gln Ala Ile Ala Asp Gly Lys Pro

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/888,320

DATE: 02/21/2002

TIME: 12:01:24

Input Set : A:\Nih413-1.app

Output Set: N:\CRF3\02212002\I888320.raw

194	65					70					75				80	
195	Ile	Leu	Glu	Tyr	Val	Lys	Ser	Thr	Ala	Ala	Met	Tyr	Gly	Ile	Asp	Arg
196					85						90				95	
197	His	Ile	Arg	Phe	His	His	Lys	Val	Ile	Ser	Ala	Asp	Trp	Ser	Thr	Ala
198				100					105					110		
199	Glu	Asn	Arg	Trp	Thr	Val	His	Ile	Gln	Ser	His	Gly	Thr	Leu	Ser	Ala
200			115						120				125			
201	Leu	Thr	Cys	Glu	Phe	Leu	Phe	Leu	Cys	Ser	Gly	Tyr	Tyr	Asn	Tyr	Asp
202		130					135					140				
203	Glu	Gly	Tyr	Ser	Pro	Arg	Phe	Ala	Gly	Ser	Glu	Asp	Phe	Val	Gly	Pro
204	145					150					155					160
205	Ile	Ile	His	Pro	Gln	His	Trp	Pro	Glu	Asp	Leu	Asp	Tyr	Asp	Ala	Lys
206					165					170				175		
207	Asn	Ile	Val	Val	Ile	Gly	Ser	Gly	Ala	Thr	Ala	Val	Thr	Leu	Val	Pro
208				180					185					190		
209	Ala	Leu	Ala	Asp	Ser	Gly	Ala	Lys	His	Val	Thr	Met	Leu	Gln	Arg	Ser
210			195					200					205			
211	Pro	Thr	Tyr	Ile	Val	Ser	Gln	Pro	Asp	Arg	Asp	Gly	Ile	Ala	Glu	Lys
212		210					215					220				
213	Leu	Asn	Arg	Trp	Leu	Pro	Glu	Thr	Met	Ala	Tyr	Thr	Ala	Val	Arg	Trp
214	225					230					235					240
215	Lys	Asn	Val	Leu	Arg	Gln	Ala	Ala	Val	Tyr	Ser	Ala	Cys	Gln	Lys	Trp
216				245						250				255		
217	Pro	Arg	Arg	Met	Arg	Lys	Met	Phe	Leu	Ser	Leu	Ile	Gln	Arg	Gln	Leu
218			260					265					270			
219	Pro	Glu	Gly	Tyr	Asp	Val	Arg	Lys	His	Phe	Gly	Pro	His	Tyr	Asn	Pro
220		275						280					285			
221	Trp	Asp	Gln	Arg	Leu	Cys	Leu	Val	Pro	Asn	Gly	Asp	Leu	Phe	Arg	Ala
222		290					295					300				
223	Ile	Arg	His	Gly	Lys	Val	Glu	Val	Val	Thr	Asp	Thr	Ile	Glu	Arg	Phe
224	305					310					315					320
225	Thr	Ala	Thr	Gly	Ile	Arg	Leu	Asn	Ser	Gly	Arg	Glu	Leu	Pro	Ala	Asp
226				325						330				335		
227	Ile	Ile	Ile	Thr	Ala	Thr	Gly	Leu	Asn	Leu	Gln	Leu	Phe	Gly	Gly	Ala
228				340					345					350		
229	Thr	Ala	Thr	Ile	Asp	Gly	Gln	Gln	Val	Asp	Ile	Thr	Thr	Thr	Met	Ala
230		355						360					365			
231	Tyr	Lys	Gly	Met	Met	Leu	Ser	Gly	Ile	Pro	Asn	Met	Ala	Tyr	Thr	Val
232		370					375					380				
233	Gly	Tyr	Thr	Asn	Ala	Ser	Trp	Thr	Leu	Lys	Ala	Asp	Leu	Val	Ser	Glu
234	385					390					395					400
235	Phe	Val	Cys	Arg	Leu	Leu	Asn	Tyr	Met	Asp	Asp	Asn	Gly	Phe	Asp	Thr
236				405						410				415		
237	Val	Val	Val	Glu	Arg	Pro	Gly	Ser	Asp	Val	Glu	Glu	Arg	Pro	Phe	Met
238				420					425					430		
239	Glu	Phe	Thr	Pro	Gly	Tyr	Val	Leu	Arg	Ser	Leu	Asp	Glu	Leu	Pro	Lys
240			435					440					445			
241	Gln	Gly	Ser	Arg	Thr	Pro	Trp	Arg	Leu	Asn	Gln	Asn	Tyr	Leu	Arg	Asp
242		450					455					460				

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/888,320

DATE: 02/21/2002
 TIME: 12:01:24

Input Set : A:\Nih413-1.app
 Output Set: N:\CRF3\02212002\I888320.raw

```

243 Ile Arg Leu Ile Arg Arg Gly Lys Ile Asp Asp Glu Gly Leu Arg Phe
244 465                               470                               475                               480
245 Ala Lys Arg Pro Ala Pro Val Gly Val
246                               485
250 <210> SEQ ID NO: 3
251 <211> LENGTH: 30
252 <212> TYPE: DNA
253 <213> ORGANISM: Artificial Sequence
255 <220> FEATURE:
256 <223> OTHER INFORMATION: Description of Artificial Sequence:EtaA PCR
257     amplification primer
259 <400> SEQUENCE: 3
260 ggggtaccga cattacgttg atagcgtgga                               30
263 <210> SEQ ID NO: 4
264 <211> LENGTH: 37
265 <212> TYPE: DNA
266 <213> ORGANISM: Artificial Sequence
268 <220> FEATURE:
269 <223> OTHER INFORMATION: Description of Artificial Sequence:EtaA PCR
270     amplification primer
272 <400> SEQUENCE: 4
273 ataagaatgc ggcgcgaacc gtcgctaaag ctaaacc                               37
276 <210> SEQ ID NO: 5
277 <211> LENGTH: 18
278 <212> TYPE: DNA
279 <213> ORGANISM: Artificial Sequence
281 <220> FEATURE:
282 <223> OTHER INFORMATION: Description of Artificial Sequence:EtaA Primer 3
283     sequencing primer, EtaA amplification primer
285 <400> SEQUENCE: 5
286 atcatccatc cgcagcac                               18
289 <210> SEQ ID NO: 6
290 <211> LENGTH: 17
291 <212> TYPE: DNA
292 <213> ORGANISM: Artificial Sequence
294 <220> FEATURE:
295 <223> OTHER INFORMATION: Description of Artificial Sequence:EtaA Primer 3
296     sequencing primer, EtaA amplification primer
298 <400> SEQUENCE: 6
299 aagctgcagg ttcaacc                               17
302 <210> SEQ ID NO: 7
303 <211> LENGTH: 17
304 <212> TYPE: DNA
305 <213> ORGANISM: Artificial Sequence
307 <220> FEATURE:
308 <223> OTHER INFORMATION: Description of Artificial Sequence:EtaA Primer 3
309     sequencing primer, EtaA amplification primer
311 <400> SEQUENCE: 7
312 gcatcgtgac gtgcttg                               17

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/888,320

DATE: 02/21/2002

TIME: 12:01:25

Input Set : A:\Nih413-1.app

Output Set: N:\CRF3\02212002\I888320.raw

L:167 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1